



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/534,661	11/10/2005	Liang He	42P14283	6515
8791 7590 12/22/2008 BLAKELY SOKOLOFF TAYLOR & ZAFMAN LLP 1279 OAKMEAD PARKWAY SUNNYVALE, CA 94085-4040				
EXAMINER				
GUERRA-ERAZO, EDGAR X				
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
12/22/2008		PAPER		

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

### Office Action Summary

**Application No.**

10/534,661

**Applicant(s)**

HE, LIANG

**Examiner**

EDGAR GUERRA-ERAZO

**Art Unit**

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 10 November 2005.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-44 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-44 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO/SE-08)  
Paper No(s)/Mail Date 06/10/2005
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_

## DETAILED ACTION

### *Introduction*

1. This office action is in response to applicant's submission filed on 11/10/2005. Claims 1-44 are currently pending and have been examined.

### *Information Disclosure Statement*

2. The Information Disclosure Statement filed on 006/10/2005 has been accepted and considered in this office action.

### *Claim Rejections - 35 USC § 102*

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless --

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. Claims 1-44 are rejected under 35 U.S.C. 102(e) as being anticipated by Kuansan et al. (EP 1 255 194). Hereinafter referred to as Kuansan.

With respect to **Claims 1, 35**, Kuansan discloses:

A method comprising: receiving user input at a client device (*user input interface 180, paragraph [0027]*); interpreting the user input to identify a selection of at least one of a plurality of web interaction modes (*the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraph [0043]*); producing a corresponding client request based in part on the user input and

the web interaction mode (*use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," paragraph [0045], Fig. 6*); and sending the client request to a server via a network (*voice recognition from audible signals transmitted by phone 80 are provided from voice browser 216 to recognition server 204, either through the network 205, or through a dedicated line 207, for example, using TCP/IP Web server 202, paragraph [0035]*).

With respect to **Claims 2, 36**, Kuansan discloses:

identifying a focused display element, the client request based in part on the identified focused display element (*Portions 282 and 284 operate similarly wherein unique recognized objects and grammars are called for each of the fields 252 and 254 and upon receipt of the recognized text is associated with each of the fields 252 and 254, paragraph [0049]*).

With respect to **Claims 3, 37**, Kuansan discloses:

sending an identifier of the identified focused display element to the server (*timeline 281 indicating when the recognition server 204 is directed to begin recognition at 283, and where the recognition server 204 detects speech at 285 and determines that speech has ended at 287, paragraph [0054]*).

With respect to **Claims 4, 38**, Kuansan discloses:

wherein the focused display element is a hyperlink (*telephony voice browser 212 receives HTML pages/scripts or the like from web server 202, paragraph [0035]*).

With respect to **Claims 5, 39**, Kuansan discloses:

wherein the focused display element is a field in a form (*the credit card information includes a field 250 for entry of the type of credit card being used, paragraph [0039]*).

With respect to **Claim 6**, Kuansan discloses:

extracting speech features from the user input, the client request based in part on the extracted speech features (*particular mode of entry, use of speech recognition with at least a display, further a screen display allowing free from selection of fields and voice recognition, paragraph [0045]*).

With respect to **Claim 7**, Kuansan discloses:

sending the extracted speech features to the server (*the Recognition element can include a "mode" attribute to distinguish the following three modes of recognition, which instruct the recognition server 204 how and when to return the results, paragraph [0053]*).

With respect to **Claim 8**, Kuansan discloses:

sending a session message to the server to initialize a connection with the server (*various attribute of the Reco element control behavior of the recognition server 204, further the attribute "initial-Timeout" 289 is the time between the start of recognition 283 and the detection of speech 285, paragraph [0055]* ).

With respect to **Claim 9**, Kuansan discloses:

wherein the session message includes an IP address of the client device (*caller's IP address, Appendix, Section 5.1 Properties*), a device type of the client device (*markup language page on the client device, paragraph [0034]*), a voice character of the user (*input data indicative of speech, DTMF, handwriting, gestures or images obtained from the user, paragraph [0034]*), a

language that the user speaks (*instruction indicating a grammar to associate with the input data entered through the client device, paragraph [0035]*), and a default recognition accuracy that the client device requests (*the return of results implies providing the “onReco” event or activating the “blind” elements as appropriate, further if the mode is unspecified, the default recognition mode can be “automatic”, paragraph [0053]*).

With respect to **Claim 10**, Kuansan discloses:

sending a transmission message to the server to exchange transmission parameters with the server (*using wireless transceiver 52 or communication interface 60, speech data is transmitted to a remote recognition server 204, further recognition results are then returned to mobile device 30 for rendering, paragraph [0016]*).

With respect to **Claim 11**, Kuansan discloses:

sending an OnFocus message to the server when a talk button is activated to notify the identifier of a focused display element (*the event “onClick” is initiated which calls or executes function “talk” in script portion 272, further this action activates a grammar used for speech recognition that is associated with the type of data generally expected in field 250, paragraph [0046]*), and the URL of current page (*telephone.Record(url, endSilence, [maxTimeout], [initialTimeout], section 5.2.6, Appendix)*).

With respect to **Claim 12**, Kuansan discloses:

sending extracted speech features to the server (*the Recognition element can include a “mode” attribute to distinguish the following three modes of recognition, which instruct the recognition server 204 how and when to return the results, paragraph [0053]*).

With respect to **Claim 13**, Kuansan discloses:

the cases to occur Unfocus message and tasks when Unfocus message occurs *(if the confidence measure is below a threshold, the “onNoReco” attribute 293 is issued, whereas if the confidence measure is above the threshold a “onNoReco” attribute 303 an the results of recognition are issued, paragraph [0057])*.

With respect to **Claim 14**, Kuansan discloses:

sending an exit message to the server to terminate a session with the server *(telephone.Hangup() instruction terminating call in progress, section 5.2.4 Appendix)*.

With respect to **Claim 15**, Kuansan discloses:

wherein a multi-modal markup language is used *(use of speech recognition in conjunction with architecture 200 and the client side markup language, further server side plug-in module 320 can generate a client side mark-up for each of the voice recognition scenarios, i.e, voice only through phone or multimodal for device 30, paragraphs [0045], [0090])*.

With respect to **Claim 16**, Kuansan discloses:

A method comprising: receiving at a server a client request from a client device via a network *(the client device can then receive input data from a user related to the field and send the data and an indication of the grammar for recognition to a recognition server, typically, located at a remote location for processing, further the remote processing devices are linked through a communications network, paragraph [0008], [0021])*; interpreting the client request to identify a selection of at least one of a plurality of web interaction modes *(providing the recognition server 204 with an indication of a grammar or language model to use during speech recognition,*

*further the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraphs [0032] [0043]), at least one web interaction mode being a speech interaction mode (use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," paragraph [0045], Fig. 6); and if the speech interaction mode is selected, receiving an identifier of a focused display element (activating a grammar used for speech recognition that is associated with the type of data generally expected in field 250, further timeline 281 indicating when the recognition server 204 is directed to begin recognition at 283, and where the recognition server 204 detects speech at 285 and determines that speech has ended at 287, paragraphs [0046], [0054]), building a correct grammar for speech recognition based on the focused display element performing speech recognition (the recognition server 204 provides an indication of a grammar or language model to use during speech recognition where upon compilation of information through recognition and any graphical user interface if used, device 30 sends the information to web server 202 for further processing and receipt of further HTML pages/scripts, paragraph [0032] ), and performing specific tasks according to the result of the speech recognition (providing particular mode of entry base based on use of speech recognition with at least a display entering for example credit card number, type of credit card, expiration date, furthermore error entry correction can be also performed paragraph [0045]).*

With respect to **Claim 17**, Kuansan discloses:

wherein the focused display element is a hyperlink (*telephony voice browser 212 receives HTML pages/scripts or the like from web server 202, paragraph [0035]).*



With respect to **Claim 18**, Kuansan discloses:

wherein the focused display element is a field in a form (*the credit card information includes a field 250 for entry of the type of credit card being used, paragraph [0039]*).

With respect to **Claim 19**, Kuansan discloses:

sending a match event to the client device via the network (*voice recognition from audible signals transmitted by phone 80 are provided from voice browser 216 to recognition server 204, either through the network 205, or through a dedicated line 207, for example, using TCP/IP Web server 202, paragraph [0035]*).

With respect to **Claim 20**, Kuansan discloses:

sending a nomatch event to the client device via the network (*voice recognition from audible signals transmitted by phone 80 are provided from voice browser 216 to recognition server 204, either through the network 205, further if the confidence measure is below a threshold, the "onNoReco" attribute 293 is issued, whereas if the confidence measure is above the threshold a "onNoReco" attribute 303 and the results of recognition are issued, paragraphs [0035], [0057]*).

With respect to **Claim 21**, Kuansan discloses:

receiving a transmission message from the client device for the exchange of transmission parameters with the client device (*executing the markup language on the client device; transmitting input data (indicative of speech, DTMF, handwriting, gestures or images obtained from the user) and an associated grammar to a recognition server remote from the client, and receiving a recognition result from the recognition server at client, paragraph [0034]*).

With respect to **Claim 22**, Kuansan discloses:

A client device comprising: a user input receiver (*user input interface 180, paragraph [0027]*); an interpreter to identify a selection of at least one of a plurality of web interaction modes from user input received by the user input receiver (*the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraph [0043]*), at least one web interaction mode being a speech interaction mode (*use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," paragraph [0045], Fig. 6*); a client request generator to generate a client request based in part on the user input and the web interaction mode, and to send the client request to a server via a network (*the remote processing devices are linked through a communications network, further voice recognition from audible signals transmitted by phone 80 are provided from voice browser 216 to recognition server 204, either through the network 205, or through a dedicated line 207, for example, using TCP/IP Web server 202, paragraphs [0021], [0035]*).

With respect to **Claim 23**, Kuansan discloses:

wherein the client request generator also identifies a focused display element, the client request based in part on the identified focused display element (*Portions 282 and 284 operate similarly wherein unique recognized objects and grammars are called for each of the fields 252 and 254 and upon receipt of the recognized text is associated with each of the fields 252 and 254, paragraph [0049]*).

With respect to **Claim 24**, Kuansan discloses:

wherein the client request generator also sends an identifier of the identified focused display element to the server *(timeline 281 indicating when the recognition server 204 is directed to begin recognition at 283, and where the recognition server 204 detects speech at 285 and determines that speech has ended at 287, paragraph [0054])*.

With respect to **Claim 25**, Kuansan discloses:

including a web interaction mode interpreter *(web enabled recognition allowing information and control on a client side to be entered, paragraph [0001])*.

With respect to **Claim 26**, Kuansan discloses:

A server apparatus comprising: a client request receiver to receive a client request from a client device via a network *(the client device can then receive input data from a user related to the field and send the data and an indication of the grammar for recognition to a recognition server, typically, located at a remote location for processing, further the remote processing devices are linked through a communications network, paragraph [0008], [0021])*; an interpreter to identify a selection of at least one of a plurality of web interaction modes from the client request received by the client request receiver *(the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraph [0043])*, at least one web interaction mode being a speech interaction mode *(use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," paragraph [0045], Fig. 6)*; a speech processor to process speech received in the

client request if the speech interaction mode is selected (*multiprocessors systems, micro-processor-based systems, network PCs operatively coupled, further timeline 281 indicating when the recognition server 204 is directed to begin recognition at 283, and where the recognition server 204 detects speech at 285 and determines that speech has ended at 287, paragraphs [0019], [0054]*), the speech processor using an identifier of a focused display element, and building a correct grammar for speech recognition based on the focused display element, the speech processor performing speech recognition, and performing specific tasks according to the result of the speech recognition (*particular mode of entry, use of speech recognition with at least a display, further a screen display allowing free form selection of fields and voice recognition, paragraph [0045]*).

With respect to **Claim 27**, Kuansan discloses:

wherein the focused display element is a hyperlink (*telephony voice browser 212 receives HTML pages/scripts or the like from web server 202, paragraph [0035]*).

With respect to **Claim 28**, Kuansan discloses:

wherein the focused display element is a field in a form (*the credit card information includes a field 250 for entry of the the type of credit card being used, paragraph [0039]*).

With respect to **Claim 29**, Kuansan discloses:

further including a web interaction mode interpreter (*the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraph [0043]*).

With respect to **Claim 30**, Kuansan discloses:

Art Unit: 2626

A multi-modal network interaction system comprising: a client device having a user input receiver (*user input interface 180, paragraph [0027]*), an client interpreter to identify a selection of at least one of a plurality of web interaction modes from user input received by the user input receiver (*the object mode provides eventing and scripting and can offer greater functionality to give the dialog author a much finer client-side control over speech interactions, paragraph [0043]*), at least one web interaction mode being a speech interaction mode, and a client request generator to generate a client request based in part on the user input and the web interaction mode, and to send the client request to a server via a network (*use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," further the client device can then receive input data from a user related to the field and send the data and an indication of the grammar for recognition to a recognition server, typically, located at a remote location for processing, further the remote processing devices are linked through a communications network, paragraphs [0008], [0021], [0045], Fig. 6*); and a server having a client request receiver to receive the client request from the client device via the network (*the client device can then receive input data from a user related to the field and send the data and an indication of the grammar for recognition to a recognition server, typically, located at a remote location for processing, further the remote processing devices are linked through a communications network, paragraph [0008], [0021]*), a server interpreter to identify a selection of at least one of a plurality of web interaction modes from the client request received by the client request receiver, at least one web interaction mode being a speech interaction mode (*the object mode provides eventing and scripting and can offer greater functionality to give the*

*dialog author a much finer client-side control over speech interactions, paragraph [0043]), and a speech processor to process speech received in the client request if the speech interaction mode is selected (particular mode of entry, use of speech recognition with at least a display, further a screen display allowing free form selection of fields and voice recognition, paragraph [0045]), the speech processor using an identifier of a focused display element (use of speech recognition in conjunction with at least a display, further this form of entry using both a screen display allowing free from selection of fields and voice recognition is called "multimodal," paragraph [0045], Fig. 6); and building a correct grammar for speech recognition based on the focused display element, the speech processor performing speech recognition, and performing specific tasks according to the result of the speech recognition (particular mode of entry, use of speech recognition with at least a display, further a screen display allowing free form selection of fields and voice recognition, paragraph [0045]).*

With respect to **Claim 31**, Kuansan discloses:

wherein the client request generator also identifies a focused display element, the client request based in part on the identified focused display element (*Portions 282 and 284 operate similarly wherein unique recognized objects and grammars are called for each of the fields 252 and 254 and upon receipt of the recognized text is associated with each of the fields 252 and 254, paragraph [0049]).*

With respect to **Claim 32**, Kuansan discloses:

wherein the client request generator also sends an identifier of the identified focused display element to the server (*timeline 281 indicating when the recognition server 204 is directed to*

*begin recognition at 283, and where the recognition server 204 detects speech at 285 and determines that speech has ended at 287, paragraph [0054]).*

With respect to **Claim 33**, Kuansan discloses:

wherein the focused display element is a hyperlink(*telephony voice browser 212 receives HTML pages/scripts or the like from web server 202, paragraph [0035]).*

With respect to **Claim 34**, Kuansan discloses:

wherein the focused display element is a field in a form (*the credit card information includes a field 250 for entry of the type of credit card being used, paragraph [0039]).*

With respect to **Claim 40**, Kuansan discloses:

A method comprising: a set of markup language (*using source document with HTML, XHTML, CHTML, XML, WML or with any other SGML-derived markup, paragraph [0042])* has been defined for applications quickly building over web by multi-modal interaction (*activating a grammar used for speech recognition that is associated with the type of data generally expected in field 250, further this type of interaction involves more than one technique of input referred to as "multimodal," , paragraph [0046]).*

With respect to **Claim 41**, Kuansan discloses:

further including: a conformance definition for the event handling of multi-modal markup language (*event handler querying the event object for data, section 2.4.3, Appendix).*

With respect to **Claim 42**, Kuansan discloses:

further including: for synchronization, two element's blocks are defined. One is sent to client and the other is kept in server *(mode employing exclusively declarative syntax, and may further be used in conjunction with declarative multimedia synchronization and coordination mechanism (synchronized markup language), further the use of speech recognition in conjunction with architecture 200 and the client side markup language, furthermore server side plug-in module 320 can generate a client side mark-up for each of the voice recognition scenarios, i.e, voice only through phone or multimodal for device 30, paragraphs [0044], [0045], [0090]).*

### ***Conclusion***

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Stawikowski et al. (U.S. Patent: 7,366,752 ) discloses a Communication system of an automation equipment based on the soap protocol.

Gao et al. (U.S. Patent: 7,152,203 ) discloses a Independent update and assembly of web page elements.

Healey et al. (U.S. Publication: 2003/0225825) discloses a Methods and systems for authoring of mixed-initiative multi-modal interactions and related browsing mechanisms.

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Edgar Guerra-Erazo whose telephone number is (571) 270-3708. The examiner can normally be reached on M-F 7:30a.m.-5:00p.m. EST. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be



reached on (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

6. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Edgar Guerra-Erazo/  
Examiner, Art Unit 2626

/Patrick N. Edouard/  
Supervisory Patent Examiner, Art Unit 2626